Biological Cybernetics manuscript No.
(will be inserted by the editor)

An algorithmic model of heading perception

(Hanada, M. (2005). An algorithmic model of heading perception. Biological Cybernetics, 92(1), 8-20.)

Mitsuhiko Hanada

Department of Cognitive and Information Sciences

Faculty of Letters, Chiba University

1-33 Yayoi-cho, Inage-ku, Chiba 263-8522, JAPAN

Current Address:

Department of Media Architecture

Future University-Hakodate

116-2 Kamedanakano-cho, Hakodate, Hokkaido, 041-8655, Japan

Abstract On the basis of Hanada and Ejima (2000, Vision Res 40:243-263) model, an algorithmic model was presented to explain psychophysical data of van den Berg and Beintema (2000, Neuron 26:747-52), which are inconsistent with vector-subtractive compensation for the rotational flow. The earlier model was modified in order not to use vector-subtractive compensation for the rotational flow. The proposed model computes the center of flow first, and then estimates self-rotation, and finally heading is recovered from the center of flow and the estimate of self-rotation. The model explains the data of van den Berg and Beintema (2000). A fusion model of rotation estimates from different sources (efferent signals, proprioceptive feedback, vestibular signals about eye and head rotation, and visual motion) was also presented.

1 Introduction

Visual motion on the retina is one of the important cues for heading. When we translate without rotation, a radial flow pattern arises on the retina and the center of the radial flow corresponds to the heading point (Fig. 1(a)). When our eye or/and body rotates, however, the center of the radial flow pattern does not correspond to the heading direction due to an additional flow by the rotation. When one translates and rotates around the vertical axis, the center of the flow pattern shifts in the direction of the rotation (Fig. 1(c)).

In pursuit eye movement, however, human observers perceive heading accurately (Royden et al. 1994; Warren and Hannon 1990). When only visual information that simulated pursuit eye movement is given to human observers, the perceived heading direction is fairly accurate under certain conditions (Li and Warren 2000; van den Berg 1993; Warren and Hannon 1990), but in many circumstances it is biased to the center of the flow pattern (e.g., Banks et al. 1996; Hanada and Ejima 2000a, c; Royden et al 1994; van den Berg 1996). For example, when observers view a visual stimulus which simulated ego-motion toward dots configured in a fronto-parallel plane while pursuing a moving dot, the perceived heading direction is consistent with the simulated heading direction. On the other hand, when observers view the same visually simulated stimulus without eye movement, they feel themselves moving to the center of the flow pattern (Royden et al. 1994). For visual stimuli which simulated ego-motion toward objects with

Fig. 1 Insert the figure about here.

different depths, the heading direction is halfway between the actual heading and the center of the flow pattern; the perceived heading is biased to the center of the flow pattern although the perceived heading does not correspond exactly to the center of flow (e.g., van den Berg 1996; Beintema and van den Berg 2000; Hanada and Ejima 2000a, c). These data suggest the following. 1) The extraretinal signals (such as efference copy or feedback signals) are used for the computation of heading in the visual system. 2) Visual information contributes to the compensation from the center of the flow pattern to the actual heading direction to some extent. The magnitude of visual compensation depends on various stimulus conditions.

It has been suggested that based on extraretinal signal, the velocity component in the retinal flow by eye movement is subtracted from the original flow, and then the heading direction was computed as the center of the residual flow. However, recently van den Berg and Beintema (2000) presented experimental results that contradict with the vector subtraction model (See also Beintema and van den Berg (2001)). When the retinal motion pattern within a small retinal aperture contains a radial pattern, precision of perceived heading is high (i.e., variance of perceived heading is small). When the pattern on the retina within the aperture becomes lamellar due to far eccentric heading or fast horizontal eye pursuit, the precision decreases. It implies that the precision of heading perception is limited by the pattern of the retinal flow, but not the pattern of flow relative to the head. On the other hand, models with vector-subtractive compensation for the rotational flow would predict that the precision depends on the

pattern of the flow relative to the head because the flow field relative to the head is recovered by the vector subtractive compensation. The result of van den Berg and Beintema (2000) implies that the vector subtraction model is implausible, and suggests that first the center of the retinal flow pattern is calculated, and then the heading direction is calculated using the center of flow and an estimate of self-rotation.

A number of models are presented for heading perception. The models recover heading from the flow field by using differential motion (Hildreth 1992; Rieger and Lawton 1985; Royden 1997), by template matching (Grossberg et al. 1999; Perrone 1992), by template matching combined with gain fields (Beintema and van den Berg 1998) or by an optimization method (Heeger and Jepson 1990, 1992; Lappe and Rauschecker 1993). We also proposed an algorithmic model of human heading perception from visual motion (Hanada and Ejima 2000a). Assuming that points whose velocity can be calculated are uniformly distributed across the visual field, our model recovers the heading direction in the following way.

1. The roll (rotation around the line of sight, see Fig. 2) is estimated as an expectation value which coincides with it, and velocity components by the roll in the flow field are removed by using the estimate.
2. The center of the radial flow pattern is estimated.
3. The yaw (rotation around the vertical axis, see Fig. 2) and pitch (rotation around the horizontal axis, see Fig. 2) are estimated by calculating expectation values which coincide with them.

4. After the rotational components in the flow field are removed, the heading direction is estimated as the center of the radial flow pattern.

5. Go back to 2 for a few iterations.

The model explains a variety of psychophysical results of heading judgement such as under-compensation for eye rotation from visual motion information alone and difference in performance for fronto-parallel planes and cloud-like stimuli (Hanada and Ejima 2000a). However, the model adopted the vector subtraction scheme. In this paper, however, we modify the model so that the vector-subtractive compensation would not be employed. The new version of the model explains the results of van den Berg and Beintema (2000).

In the earlier work (Hanada and Ejima 2000a), we did not present an explicit way for combining the visual information with extraretinal information about self-rotation. In this paper, we will propose a fusion model for extraretinal estimates of self-rotation and a visual estimate of self-rotation. Furthermore, we will discuss some phenomena about heading perception that may be explained by the present model in General discussion.

2 Estimation method for heading computation in the visual system

In this section, we propose a simple algorithmic model for heading perception. This model is based on Hanada and Ejima (2000a) model of heading judgement. Since derivations of equations that will be

presented below is straightforward from mathematical analyses in the earlier paper, we omit the details of the derivations.

2.1 Velocity flow by ego-motion

We assume that an observer (or camera) translates forward in the 3-D rigid environment while rotating and the rotation rate is low. We make use of essentially the same notation as Longuet-Higgins and Prazdny (1980). We use a coordinate system that is fixed with respect to an observer. The translation of the observer in the rigid environment is expressed in terms of translation along three orthogonal directions, which we denote by the vector (U, V, W). U, V and W show translation along the X-axis, Y-axis and Z-axis respectively (Fig. 2). The Z-axis is directed along the optical axis, and the X-axis and Y-axis are horizontal and vertical respectively. The rotation of the observer is expressed in terms of rotation around three orthogonal axes, which we express by the vector (A, B, C). A, B and C, which show rotation around the X-axis, the Y-axis and the Z-axis, respectively (Fig. 2). The 3-D velocity of a point, (X, Y, Z) is given by

$$\dot{X} = -U - BZ + CY \tag{1}$$

$$\dot{Y} = -V - CX + AZ \tag{2}$$

$$\dot{Z} = -W - AY + BX \tag{3}$$

Fig. 2 Insert the figure about here.

where $(\dot{X}, \dot{Y}, \dot{Z}) \equiv (dX/dt, dY/dt, dZ/dt)$ (Longuet-Higgins and Prazdny 1980). If we consider perspective projection of the velocity onto the image plane $Z = 1$ for the projection, point P on the image (x, y) is given by

$$x = X/Z \tag{4}$$

$$y = Y/Z \tag{5}$$

The projected velocity $(u, v) \equiv (\dot{x}, \dot{y}) \equiv (dx/dt, dy/dt)$ in the image plane is given by (Longuet-Higgins and Prazdny 1980)

$$u = \frac{-U + xW}{Z} - B + Cy + Axy - Bx^2 \tag{6}$$

$$v = \frac{-V + yW}{Z} - Cx + A + Ay^2 - Bxy \tag{7}$$

Let (x_i, y_i), (u_i, v_i) and Z_i be the projected position, the velocity and the depth of the i- th sampling point, respectively. The visual system must recover heading from (x_i, y_i) and (u_i, v_i) $(i = 1, \ldots, N)$. In principle, translation (U, V, W) can not be recovered from the velocities, but only the ratio of U, V and W $(U : V : W)$ can be recovered (i.e., only the heading direction can be recovered).

2.2 Algorithm

The algorithmic model assumes the following to recover heading.

Fig. 3 Insert the figure about here.

1. Self-rotation is slow.

2. There is a large number of sampling points which are randomly distributed in a large visual field.

3. There are large depth variations for the scene.

The algorithm for the model is very simple; first the center of flow is estimated, second the self-rotation is estimated. Based on the center of flow and the estimated self-rotation, the heading direction is estimated.

2.2.1 Estimation of the center of flow First we calculate the center of flow. The center of flow is the point that minimizes the square sum of the distance between the point and the line passing through the velocity flow vector (Fig. 3). (In the earlier papers (Hanada and Ejima 2000a, b) we called it the center of outflow. In this paper, however, we use the center of flow because Lappe and Rauschecker (1994) and Beintema and van den Berg (2000) used the center of flow to express essentially the same concept.) The point is easily calculated using a linear least-square method.

2.2.2 Coordinates transform We transform the coordinate system in the image so that the center of flow is the origin (Fig. 4). Let (X', Y', Z') denote the transformed coordinates of a 3D point which is represented as (X, Y, Z) in the original coordinates. Let (U', V', W') and (A', B', C') be translation and rotation in the new coordinates, respectively. Let $(x', y') = (X'/Z', Y'/Z')$ (the image position projected to the new projection plane $Z' = 1$ in the new coordinates), and let (u', v') be the image velocity on

Fig. 4 Insert the figure about here.

the new projection plane. Let $\mathbf{P} = (X, Y, Z)^t$, $\mathbf{P}' = (X', Y', Z')^t$, $\mathbf{T} = (U, V, W)^t$, $\mathbf{T}' = (U', V', W')^t$, $\mathbf{R} = (A, B, C)^t$, $\mathbf{R}' = (A', B', C')^t$, $\mathbf{p} = (x, y, 1)^t$, $\mathbf{p}' = (x', y', 1)^t$, $\mathbf{t} = (u, v, 0)^t$ and $\mathbf{t}' = (u', v', 0)^t$ where ^{t} denotes transpose. The whole transform equations are given by

$$\mathbf{P}' = \mathbf{H}^t\mathbf{P}, \quad \mathbf{T}' = \mathbf{H}^t, \quad \mathbf{R}' = \mathbf{H}^t\mathbf{R}$$

$$\mathbf{p}' = m(\mathbf{H}^t\mathbf{p}), \qquad \mathbf{t}' = l(\mathbf{H}^t\mathbf{p}, \mathbf{H}^t\mathbf{v}) \tag{8}$$

where $m((a, b, c)^t)$ denotes $(a/c, b/c, 1)$, $\quad l((a, b, c)^t, (e, f, g)^t)$ denotes $((e - ag)/c, (f - bg)/c, 0)$, and $\mathbf{H}$ is a 3×3 orthogonal matrix from the coordinates $(X, Y, Z)^t$ to the new coordinates $(X', Y', Z')^t$. Let the center of flow be (x_c, y_c). Let azimuth and elevation of the center of flow be ϑ and φ, respectively. The transform matrix $\mathbf{H}$ is given by

$$\mathbf{H} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & cos\,\varphi & sin\,\varphi \\ 0 & -sin\,\varphi & cos\,\varphi \end{pmatrix} \begin{pmatrix} cos\,\vartheta & 0 & sin\,\vartheta \\ 0 & 1 & 0 \\ -sin\,\vartheta & 0 & cos\,\vartheta \end{pmatrix} \tag{9}$$

where

$$(sin\,\vartheta,\ cos\,\vartheta) = (x_c/\sqrt{1 + x_c^2},\ \ 1/\sqrt{1 + x_c^2}) \tag{10}$$

$$(sin\,\varphi,\ cos\,\varphi) = (y_c/\sqrt{1 + x_c^2 + y_c^2},\ \ \sqrt{1 + x_c^2}/\sqrt{1 + x_c^2 + y_c^2}) \tag{11}$$

2.2.3 From the center of flow to the heading direction Hanada and Ejima (2000a) pointed out that for no rotation around the Z' axis (i.e., C'=0), one can think that the observer is tracking a virtual point whose

projection is the center of flow, and whose depth is approximately the average of other sampling-points' depth (that is, tracking a point (0, 0, Z'_0) in the transformed coordinates where Z'_0 is average depth for all the points (x'_i, y'_i) $(i = 1, \cdots, N)$). This also approximately holds with some rotation around the Z' axis. (Since the derivation is directly obtained from the mathematical analysis of Hanada and Ejima (2000a), we omit it.) Therefore we obtain (Hanada and Ejima 2000a)

$$0 \approx \frac{-U'}{Z'_0} - B', \quad 0 \approx \frac{-V'}{Z'_0} + A' \tag{12}$$

Time to contact (collision) of the virtual point (0, 0, Z'_0) is ($\equiv Z'_0/W'$), and is approximately average time to contact of all the sampling points.

$$\hat{\tau} = \left\langle \frac{x_i'^2 + y_i'^2}{u'_i x'_i + v'_i y'_i} \right\rangle \approx Z'_0/W' \tag{13}$$

where a hat above a symbol indicates the estimate of a value represented by the symbol, and the brackets denote ensemble averaging over dots in a symmetric region around the center of flow (but not the image origin). For the averaging, we should exclude the sampling points such that the denominator of the value inside the brackets is not too small. (If $u'_i x_i + v'_i y'_i \approx 0$, then the dot should be ignored for the calculation of the time to contact.) From (12) and (13), we obtain

$$0 \approx -U'/(\hat{\tau}W') - B', \quad 0 \approx -V'/(\hat{\tau}W') + A' \tag{14}$$

From these equations, the following equation is derived.

$$(U'/W',\ V'/W') \approx (-B'\hat{\tau},\ \ A'\hat{\tau}) \tag{15}$$

The equation shows the shift from the center of flow to the heading point on the new projection plane. If we know rotation (A', B', C'), therefore, we can compute the heading direction. Hence the next problem is estimation of rotation.

2.3 Estimation of rotation

2.3.1 Efference copy, proprioceptive feedback and vestibular signals Several estimates of rotation are obtained from different extraretinal sources of information. We can use information of efference copy, proprioceptive feedback from muscles of eyes and a neck, and information of the vestibular system to estimate rotation (A', B', C'). We left the concrete mechanisms of the estimation black boxes in this paper.

2.3.2 Visual motion We presented a method for estimating rotation from only visual motion information in the earlier work. The method estimates rotation parameters calculating expectation values which coin-

cide with the parameters, assuming that points whose velocity can be calculated are uniformly distributed across the visual field. Rotation around the line of sight is estimated as

$$\hat{C}' = \begin{cases} \left\langle \dfrac{u'_i y'_i - v'_i x'_i}{x'^2_i + y'^2_i} \right\rangle & \text{for general situations} \\ \left\langle \dfrac{v'_i}{x'_i} \right\rangle & \text{for ego-motion on the ground} \end{cases} \tag{16}$$

Using this estimate of C', we estimate A' and B' as follows. (For the derivations, see Hanada and Ejima (2000a)).

$$\hat{A}' = \left\langle \omega_i \frac{(v'_i + x'_i \hat{C}') - (u'_i - y'_i \hat{C}') y'_i / x'_i}{1 - \hat{\tau}/\hat{\tau}_i} \right\rangle \tag{17}$$

$$\hat{B}' = \left\langle \xi_i \frac{(v'_i + x'_i \hat{C}') x'_i / y'_i - (u'_i - y'_i \hat{C}')}{1 - \hat{\tau}/\hat{\tau}_i} \right\rangle \tag{18}$$

where

$$\hat{\tau}_i = \frac{x'^2_i + y'^2_i}{u'_i x'_i + v'_i y'_i} \tag{19}$$

The value indicate an estimate of time to contact for the i-th sampling point. Values ω_i and ψ_i are weights for summation, which satisfies $\sum \omega_i = 1$ and $\sum \psi_i = 1$. One simple way for weighting are to set all the weights to a single value (i.e., simple averaging). Instead, Hanada and Ejima (2000a) suggested another way for weighting as follows.

$$\omega_i = \frac{1}{\Omega} \exp\left(-\left(\frac{y^2/x^2}{\sigma^2}\right)\right), \quad \xi_i = \frac{1}{\Xi} \exp\left(-\left(\frac{x^2/y^2}{\sigma^2}\right)\right) \tag{20}$$

where Ω and Ξ are constants to make summation of the weights 1. We suggested that σ be about 0.3 from results of psychophysical experiments (Hanada and Ejima 2000a). For the estimation of A and B, velocities due to C are subtracted from the flow based on the estimate of C, and in fact vector subtraction is performed at this stage. However, the vector subtraction does not contradict with the empirical data of van den Berg and Beintema (2000) because their data only suggested that velocity components by rotation are not subtracted for the calculation of the center of flow. Estimation of rotation around the X' and Y' axes may be performed through subtraction of velocity vectors due to rotation around the Z' axis.

2.3.3 Fusion of rotation estimates The rotation estimates from the different sources should be fused. The weak fusion model is successful for explanation of cue integration about depth (Young et al. 1993; Landy et al. 1995). In the weak fusion model, the visual system weighs depth cues according to the reliability and linearly combines the depth cues using the weights. However, the weak fusion model seems insufficient to explain the data of heading judgement. For example, cue integration about self-rotation was experimentally examined, and it was shown that cue integration is not linear weighting of cues (Crowell

et al. 1998). A new model may be needed to explain how the cues of self-rotation are used for heading judgement. A fusion model will be presented later.

It should be noted that fusion of a variety of rotation estimates can be performed because the vector subtraction for recovering the head-centric flow is not employed and the visual estimate of rotation is processed in the same way as the other estimates from different sources.

2.4 From the new coordinates to the original ones

The heading direction corresponds to the direction of the vector $S = (U'/W', V'/W', 1)$. The heading vector and the rotation parameters can be converted to the representations in the original coordinates as follows.

$$\mathbf{S} = \mathbf{H}\mathbf{S}', \quad \mathbf{R} = \mathbf{H}\mathbf{R}' \tag{21}$$

2.5 Summary of the algorithmic model

We summarize the algorithmic model presented here.

1. We find the center of flow.
2. We transform the coordinates so that the center of flow is projected to the image origin using (8).
3. We calculate $\hat{\tau}$ using (13).

4. We estimate C', A' and B' by extraretinal and vestibular information, and by (16), (17) and (18), and fuse the estimates in some ways.

5. We estimate (U'/W', V'/W') using (15).

6. We convert the heading direction and the rotation parameters to those in the original coordinates using (21).

3 Difference between the earlier and present models

For the earlier version of the model, a method for improving the estimate of heading by iterative computation was proposed (Hanada and Ejima 2000a); rotation is recovered and the rotational velocity components are removed, and then heading is recovered. Then the remains of the self-rotation is recovered from the residual flow, and the rotational flow is again removed. The procedure was repeated several times. For the earlier and present algorithms, it is assumed that rotation is slow. Since rotation to be estimated decreases with the iteration, accuracy of the estimate improves with the iteration. For the present model, however, we do not adopt the iterative computation since vector subtraction is essential for the iteration. Although the iteration technique improves accuracy of heading computation, it seems unnecessary for the explanation of human heading judgement. Human heading perception from visual information alone has some bias toward the center of flow for self-rotation and the accuracy is not so high. Furthermore, simulation results of Hanada and Ejima (2000a) showed that the non-iterative model can

explain many psychophysical data of heading perception. Hence, we abandon the iterative computation. Consequently we also modified the following points.

1. In the earlier model, first roll (C, rotation around the line of axis) is estimated and velocity components due to roll in the flow field are removed. Since the idea of vector subtraction is abandoned for this model, all components of rotation are estimated in a later stage, and then the heading direction is recovered without the vector subtraction.
2. Time to contact of each point and average time to contact are calculated in a slightly different way. The earlier model estimates time to contact for a point in the following way; $\hat{\tau}_i = (x_i'^2 + y_i'^2)/(u_i'^2 + v_i'^2)$. The estimate is not so accurate when there is fast rotation around the Z' axis. The earlier model removes velocity components due to C first, and then assuming $C \approx 0$, time to contact is estimated. Since rotation is estimated at a later stage in the present model, we must estimate time to contact from the flow including velocity components due to rotation around the Z' axis. For the present model, we use only radial velocity components for the calculation of time to contact. Note that the estimate by Eq. (19) is almost invariant to the rotational flow such as the lamellar flow caused by A and B or the circular flow due to C.

Since the modifications do not change essence of the computation (especially in the case of no roll), the demonstrations by the simulations in the earlier paper are also replicable for the modified model.

4 Fusion of estimates of self-rotation from different sources

The visual system uses extraretinal information about eye rotation to compensate for effects of eye movement on the retinal motion pattern. Earlier researches indicate that this compensation is nearly complete (e.g., Banks et al. 1996; Royden et al. 1994). Also, the visual system uses extraretinal information about head turn and the compensation is nearly perfect. There are three possible sources of extraretinal information for head turn: efferent information about the motor commands to neck muscles, proprioceptive information from neck muscles and vestibular canal information about head rotation. Crowell et al. (1998) showed that at least two out of the three sources of information was required for the compensation for head turns. The visual system may adopt median of the rotation estimates from the three sources as the fused estimate.

When observers move toward a fronto-parallel plane while pursuing a moving dot, they can compensate for the rotational flow due to eye movement nearly perfectly. However, they cannot compensate for the rotational flow at all when the flow is simulated only visually (Royden et al. 1994; Warren and Hannon 1990). It indicates that the visual estimate of rotation is zero for a fronto-parallel plane and that human observers can compensate for the rotational flow due to eye movement only from extraretinal information. Under some visually rich conditions, however, human observers can compensate for the rotational flow only from visual information without extraretinal information (e.g., Li and Warren 2000; Warren and

Hannon 1990). It seems that if either visual or extraretinal estimate of rotation is correct, human observers can compensate for the rotational flow. The visual system may fuse visual and extraretinal estimates by adopting maximum of them as the final estimate.

On the basis of the analysis, we present a fusion model for rotation estimation.

4.1 Estimation from separate sources

Let $\mathbf{R}$ be rotation (A, B, C). First rotation is estimated from different sources. The estimates are denoted as follows.

1. $\hat{\mathbf{R}}_{e,e}$: estimate of eye movement from efference copy of eye movement
2. $\hat{\mathbf{R}}_{e,p}$: estimate of eye movement from proprioceptive sensors of eye muscles
3. $\hat{\mathbf{R}}_{b,e}$: estimate of head turn + body turn from efference copy
4. $\hat{\mathbf{R}}_{b,p}$: estimate of head turn + body turn from proprioceptive information from neck and body muscles
5. $\hat{\mathbf{R}}_{b,v}$: estimate of head turn + body turn from vestibular information
6. $\hat{\mathbf{R}}_{m}$: estimate of eye rotation + head turn + body turn from visual motion

Assuming that body turn and head turn are processed in the same way, information from efference copies and proprioceptive sensors about head turn and body turn are summed up and dealt with altogether for this fusion model.

4.2 Fusion of estimates

Here we propose a fusion model of the above estimates from different sources.

1. Fusion of eye-rotation estimates

   Select from $\hat{\mathbf{R}}_{e,e}$ and $\hat{\mathbf{R}}_{e,p}$ the one that has the minimum or median absolute value and set it to $\hat{\mathbf{R}}_e$, which represents an estimate of eye rotation.

2. Fusion of head turn + body turn estimates

   Select from $\hat{\mathbf{R}}_{b,v}$, $\hat{\mathbf{R}}_{b,e}$ and $\hat{\mathbf{R}}_{b,p}$ the one that has the median absolute value and set it to $\hat{\mathbf{R}}_b$, which represents an estimate of head turn + body turn.

3. Estimate of rotation from extraretinal information

   The estimate of rotation from extraretinal information ($\hat{\mathbf{R}}_x$) is the sum of the estimates of eye rotation and body-head turn; $\hat{\mathbf{R}}_x = \hat{\mathbf{R}}_e + \hat{\mathbf{R}}_b$

4. Estimate of total rotation

   Select from $\hat{\mathbf{R}}_x$ and $\hat{\mathbf{R}}_m$ the one that has the maximum absolute value and set it to $\hat{\mathbf{R}}$, which represents an estimate of total rotation.

4.3 Comparison of the model's prediction and human performance

4.3.1 Warren and Hannon (1990). Warren and Hannon (1990) compared performance of heading judgement under two conditions: (a) the observer tracked a moving point, introducing a rotational component

of motion ('moving' condition). (b) the observer maintained stationary fixation while the display contained both translational and rotational components of motion ('simulated' condition). The same flow pattern appeared on the retina for the conditions (a) and (b), though the rotation information could be derived from extraretinal sources for (a). For the case of movement toward a cloud of random dots or a ground plane, observers performed heading judgement accurately and there was essentially no difference in performance between the conditions. In other words, nearly perfect compensation was observed for (a) and (b). However, when simulating translation toward a fronto-parallel plane, human observers perceived heading accurately for the 'moving' condition, but did not at all for the 'simulated' condition; nearly perfect compensation for (a), but little compensation for (b).

First we consider the case of a fronto-parallel plane. The proposed model cannot estimate rotation from the flow for a fronto-parallel plane (Hanada and Ejima, 2000a), and it may be assumed that $\hat{\mathbf{R}}_m = 0$. For the 'moving' condition (a), we may assume that $\hat{\mathbf{R}}_{e,e} = \hat{\mathbf{R}}_{e,p} = \mathbf{R}$. For the 'simulated' condition (b), however, $\hat{\mathbf{R}}_{e,e}$ and $\hat{\mathbf{R}}_{e,p}$ would be zero. Since there was no body and head turn in their experiment, $\hat{\mathbf{R}}_{b,e}$, $\hat{\mathbf{R}}_{b,p}$ and $\hat{\mathbf{R}}_{b,v}$ would be zero. $\hat{\mathbf{R}}$ (the total estimate of rotation) becomes $\mathbf{R}$ for (a), but zero for (a). Hence perfect compensation should occur for (a) and no compensation should occur for (b).

Next we consider the case of a cloud or ground plane. For their experiment, rotation was less than 1.5 deg/s and very small, and there was enough depth variation for rotation estimation. Hence, we may assume that $\hat{\mathbf{R}}_m = \mathbf{R}$. (Very accurate heading perception justifies the assumption.) The other estimates

are the same as in the case of a fronto-parallel plane. Thus, $\hat{\mathbf{R}}$ are $\mathbf{R}$ for both (a) and (b). Hence perfect compensation should occur for (a) and (b).

Thus, the proposed model explains the data of Warren and Hannon (1990) fairly well.

4.3.2 Crowell et al. (1998). The present model focuses on perception of the instantaneous heading, while observers judged the destination point in future in the experiments of Crowell et al. (1998). The under-compensation for the rotational flow may be not due to misperception of instantaneous heading, but due to curved path perception (Royden 1994). However, we assume here that the under-compensation in the experiments of Crowell et al. (1998) is caused by, or at least linearly related to misperception of instantaneous heading.

Crowell et al. (1998) examined visual self-motion in the following conditions. *Eye pursuit*; the observers tracked the target by turning the eye in the orbit with the head stationary. In this condition, proprioceptive information and information of efference copy about eye rotation was present, but vestibular information was absent. *Active head pursuit*; observers pursued a moving target by turning the head with the eye fixed in the orbit. In this condition, there was proprioceptive information, vestibular information and information of efference copy about head turn. *Passive head pursuit*; the observer's head was turned by a motorized chin/head rest. In this condition, vestibular and proprioceptive information about head turn was present, but information of efference copy was absent. *Active head stabilization*; the body was turned

while the observer actively counter-rotated the head to hold it fixed in space. In this condition, there were proprioceptive information and information of efference copy about head turn, but there was no vestibular information. *Passive head stabilization*; the body rotated while the head was held fixed in space. In this condition, proprioceptive information about head turn was present, but vestibular information and information of efference copy was absent. *Full body rotation*; the observer's entire body was rotated by a motorized chair while the head was held fixed with respect to the body. In this condition, vestibular information about head turn was present, but proprioceptive information and information of efference copy was absent. *Simulated pursuit*; the observer held the gaze fixed while the display simulated a gaze shift. In this condition, there was no extraretinal information about rotation.

They found that errors in the perceived destination direction are proportionate to the gaze rotation rate. The slope reflects the level of performance. No slope means perfect compensation, and the large slope indicates poor compensation. They found that in the 'simulated-pursuit' conditions, errors were about 20 deg at a rotation rate of 20 deg. The slope of the function of path error vs. gaze rotation rate was fairly sharp, which indicates that the visual compensation was poor. They used that performance as benchmark in the absence of extraretinal compensation. Let the slope in this condition be M_{sim}. They defined compensation index CI [%] as $100 \times (1 - M/M_{sim})$ [%] where M is the slope in the condition in question. This index quantifies the effectiveness of a given extraretinal cue for accurate self-motion judgement.

Crowell et al. (1998) reported that mean CI was 90% during eye pursuit, and 94% during active head pursuit, 4% during body rotation, 21% during passive head stabilization, 68% during active head stabilization, and 78% during passive head pursuit. The results indicate that when two out of the sources of information about head turn were available, there was considerable compensation for rotation.

From Eq. (15), we can easily show that the proposed model predicts CI for certain $\mathbf{R}$ as follows.

$$CI = 100 \times \left(1 - \frac{1 - |\hat{\mathbf{R}}|/|\mathbf{R}|}{1 - |\hat{\mathbf{R}}_m|/|\mathbf{R}|}\right) = \frac{|\hat{\mathbf{R}}| - |\hat{\mathbf{R}}_m|}{|\mathbf{R}| - |\hat{\mathbf{R}}_m|} \quad [\%] \tag{22}$$

We assume that the proprioceptive systems for neck and eye and the vestibular systems for head turn can estimate rotation virtually perfectly, and estimates of head and eye rotation from efference copy are perfectly accurate. From these assumptions, we can easily computed the predicted CI, and the predicted CI is independent of $\mathbf{R}$ and $\hat{\mathbf{R}}_m$ (the visual estimate of rotation). The predicted CIs are shown in Fig. 5 with the average CI for human observers. For fully-body rotation and passive head stabilization, the CIs for the model are zero, and those for human observers are also very small. For the other conditions, the CIs are 100 %, and the CIs for human observers also are fairly high (greater than 68 %). The predicted CIs are qualitatively consistent with human data. However, difference in CI between human observers and the model is rather conspicuous for active head stabilization and passive head pursuit. However, there were large individual differences of the CIs for these conditions in their experiments. Moreover, the CIs in these conditions were unstable for an individual (Crowell et al. 1998). Hence it may be impossible

Fig. 5 Insert the figure about here.

to explain the whole data completely in these conditions in principle. Hence, the proposed fusion model would be good enough as a practical approximation. It should be noted that the CIs for an observer in these conditions were almost the same as those in the conditions of active head pursuit and eye rotation (Crowell et al. 1998). The CIs for the observer were very similar to the predicted CIs.

4.3.3 Banks et al. (1996). Banks et al. (1996) examined heading judgement in the presence of rotation consisting of mixtures of executed and simulated eye movement. The proportion of simulated eye rotation was varied. They found that errors in judged heading depend mainly on the magnitude of simulated eye rotation, but not on the total rotation rate. We next show that the proposed model explains the result well.

In the 'simulated' conditions in Experiment 1 of Banks et al. (1996) (i.e., when the proportion of simulated eye rotation was zero), an error in heading judgement was proportionate to a simulated rotation rate. For a simulated rotation rate of 7.5 deg/s, it was between 10 deg to 15 deg. In that experiment, average time to contact for their stimuli was 1.7 [s] at the beginning of the presentation, and the distance between the heading and the center of flow was about 13 deg. It implies that their observers responded to points around the center of flow as the heading point and the rotational flow were little compensated for. The larger error may be due to the rather small number of dots (64 dots at the beginning of the

presentation and much less at the end). Also, it might be due to perception of curved path (Royden, 1994). Another possible reason is that fairly high simulated rotation rate were used in their experiments than those of Warren and Hannon (1990).

Hence, we may assume that the visual estimate of rotation is zero; $\hat{\mathbf{R}}_m = 0$. Hence, $\hat{\mathbf{R}} = \hat{\mathbf{R}}_e$. It is likely that the estimate of actual eye rotation is very accurate. Hence, we may assume that $\hat{\mathbf{R}}_e$ is actual eye rotation. From Eq. (15), the error in perceived heading was $(|\mathbf{R} - \hat{\mathbf{R}}|)\tau$. Note that $|\mathbf{R} - \hat{\mathbf{R}}|$ equals the simulated rotation rate. Hence, the model predicts the heading error is proportionate to the simulated rotation rate. The slope of the function of the heading error vs. the simulated rotation rate should be average time to contact τ regardless of the total rotation rate. The prediction agrees on their results very well (See Fig. 4 of Banks et al. al (1996)).

4.4 Discssion

Lappe (1998) proposed a method for combining the flow and extraretinal signals for their neural-network model of heading perception. The method is essentially equivalent to the approach that after subtracting out the rotation using the extraretinal signal, heading is recovered from the remaining flow. Beintema and van den Berg (2000) proposed multiplicative inhibition to head-centered template cells, whose magnitudes are based on the degree of conflict between the extraretinal signals and visual estimation of rotation. The

'maximum' rule for the fusion of extraretinal and visual estimates of rotation is much simpler than the rules used by those studies, but it can explain human data fairly well.

We used different assumptions about visual estimation of rotation; nearly perfect estimation for Warren and Hannon (1990) and zero estimation for Banks et al. (1996). However, they were not arbitrary assumptions. The accuracy of visual estimation was based on human performance in the 'simulated' condition of their experiments. Degree of visual compensation for the rotational flow was varied across the studies. The possible reasons for the difference of visual compensation will be shown in General discussion.

Banks et al. (1996) restricted eye pursuit to one in the same direction as the simulated rotation. van den Berg et al. (2001) examined path perception during eye pursuit in the opposite direction as the simulated rotation as well as during pursuit in the same direction. They measured perceived path, and derived the perceived destination point, the perceived curvature and the perceived instantaneous heading from the perceived path. They replicated the results of Banks et al. (1996) about perceived destination points for eye pursuit in the same direction of simulated rotation. However, they reported complicated interaction of the perceived heading (or destination), the simulated rotation rate and the direction of actual eye rotation. Also, there were large individual differences in the perceived heading. The proposed model may not explain their data, especially those for eye pursuit in the opposite direction to the simulated rotation. More complicated models for fusion of visual and extraretinal rotation estimates are needed to explain their data. However, large individual differences make the modeling very difficult.

We leave the fusion model for eye pursuit in the opposite direction to the simulated rotation for future work.

Although the model is consistent with many psychophysical data, we do not have clear reasons why conservative estimates such as median of different kinds of estimates are used for extraretinal estimation of rotation and why maximum of the visual estimate and the extraretinal estimate is used for the total estimation. (Maybe because humans rely on visual information more than on vestibular and proprioceptive information as input.) In this sense, the fusion model is not computational. However, since prediction of the model is unambiguous, we can test the model experimentally. The fusion model would give a framework for further computational and empirical studies.

5 Simulation of the psychophysical experiment of van den Berg and Beintema (2000)

van den Berg and Beintema (2000) measured precision of perceived heading varying the heading direction and the eye rotation rate, and they showed that the precision depended on the retinal flow, but not on the head-centric flow, and the precision was highest when the center of the retinal flow pattern was visible within the stimulus aperture. We performed simulations of their experiment using the present model.

The input to the present model was velocities of dots within an aperture (diameter, 10 deg) at the final frame of stimuli used in their experiment. The simulated environment was as follows. The stimulus was random dots configured in a fronto-parallel plane. The simulated speed of translation was varied

between 1.0 and 2.0 m/s. Depth of the simulated plane was depth at the final frame in their experiment, that is, $9 - W \times D$, where W is the speed of simulated self-motion and D was the presentation duration, which was 0.5 s. (The simulated initial depth was 9 m in their experiment.) The number of dots was not specified explicitly in their paper, although they stated that in the final frame more than 20 dots were visible in any conditions. Hence, we used velocities of 40 dots as input. We computed the velocities of the dots. The speed noise was added to the velocities. The speed noise was Gaussian noise whose standard deviation was 10 % of the velocity of the dot. The direction was also perturbed by Gaussian noise whose SD was 2.5 deg. (The magnitude of the noise was selected according to the data of the direction and speed discrimination thresholds by De Bruyn and Orban (1988).) In the experiment of van den berg and Beintema (2000), the fixation point was fixed or moved at a rate of -3.0 or 3.0 deg/s to induce pursuit. Hence the simulated rate of eye rotation was -3.0, 0 or 3.0 deg/s. For the stimulus conditions, 3.0 deg/sec eye rotation displaces the center of the radial pattern on the retina (the retinal focus) about 18 deg at the beginning of the stimulus presentation and about 16.5 deg at the end of it. In this simulation, we assumed that the rate of eye rotation was estimated by extraretinal information alone. We used actual rotation rates plus Gaussian noise whose SD was 0.1 deg/s as the estimates of rotation. We assumed that the error of pointing was Gaussian distributed with SD of 1.0 deg, and this SD was added to SD of the heading estimates of the present model to obtain SD of the model's responses. The center of flow was estimated as a concurrent point of lines whose directions were the velocities of dots by the method

Fig. 6 Insert the figure about here.

of Kanatani (1991). Strictly speaking, the concurrent point is not the center of flow of our definition. For a lamellar flow, however, small noise disrupts the computation of the center of flow of our definition because the intersection points of some flow lines are located at the opposite side of the actual center of flow as shown in Fig. 6. This problem can be avoided by using homogeneous coordinates of projective geometry because homogeneous coordinates include infinite points and parallel lines are intersected at an infinite point. Since the experiment of van den Berg and Beintema (2000) examined difference in human performance of heading judgement between radial patterns and lamellar ones and lamellar flows occurred in many conditions, we used the method of Kanatani (1991).

We conducted 1000 trials to obtain SD for each condition. The SDs obtained by the simulation are shown in Fig. 7(a). The SD of the heading direction estimated by the present model was a U-shaped curve with a minimum. For 0 deg/sec eye rotation, the SD was lowest at simulated heading of $\pm$ 4.0 deg. Namely, SD of perceived heading were minimal when the center of the radial flow pattern was visible within the stimulus aperture. It is because small noise to the nearly parallel flow vectors shifts the intersections formed by the flow lines while noise to the nearly orthogonal flow vectors shift them to a smaller extent. For eye rotations of $\pm$ 3.0 deg/sec, the minima of SDs were located at heading directions of about $\pm$ 20 deg. As a function of the retinal focus (shifting the curves for pursuit by 18 deg in the pursuit conditions as in van den Berg and Beintema (2000)), the SD curves match better and the highest

Fig. 7 Insert the figure about here.

precision was found when the center of the radial pattern was visible on the retina. However, the SD curves did not match precisely for the three eye rotations. The SD of heading estimates for eye rotation of 3.0 deg/sec (resp. -3.0 deg/sec) does not match the SD for no eye rotation in the positive (resp. negative) retinal focus. These characteristics of the SD curves were also found in the SD curves for human observers (van den Berg and Beintema 2000). Although the SD curves as a function of the heading direction were slightly flatter than the curves of human observers, the shape was similar to that in their paper, and the present model is qualitatively consistent with human data.

For human observers, the SD curves as a function of the retinal focus mismatch for different eye rotations, and van den Berg and Beintema (2000) ascribed it to lower pursuit gain. We performed the same simulation using the actually simulated time to contact for the heading computation of the present model and found that the SD curves matched very well. Therefore, the reason for the mismatch in this simulation of the present model is imprecision of the estimate of time to contact. The mismatch of the SD curves for human observers may be ascribed to this reason, and not to lower pursuit gain.

5.1 Discussion

In the case of the front-parallel plane and real eye movement, the heading computation used by van den Berg and Beintema (2000) and Beintema and van den Berg (2001) as the working hypothesis was

essentially the same as the computation of Hanada and Ejima (2000a) model without iteration (and the present model), though they did not mention it. As a result, their analysis was similar to the analysis through this simulation. However, it is not identical to this analysis. First in their analyses the center of flow was estimated assuming horizontal translation. On the other hand, we do not use the assumption. Moreover, they did not discuss the problem for the computation of the center of flow for lamellar flows described above. Second van den Berg and Beintema (2000) do not discuss accuracy and precision of estimation of time to contact. When estimation of the center of flow is inaccurate and/or imprecise, an estimate of time to contact calculated from the center of flow is inaccurate and/or imprecise. For lamellar flow patterns with noise, the estimation of the center of flow is imprecise. It leads to imprecision of the estimation of time to contact. If heading is computed by (15), imprecision of heading estimation is caused not only by imprecision of the estimation of the center of flow but also by imprecision of the estimation of time to contact. When rotation is fast or when the estimation of rotation is imprecise, estimation of heading is largely affected by the estimate of time to contact since heading is computed by the product of time to contact and the rotation estimate. Hence, time to contact is not a negligible factor for their experiments. Thus, time to contact was estimated from the flow in this simulation. The present simulation showed that imprecision of estimation of time to contact leads to the mismatch of the SD curves as a function of the retinal focus. Hence the present analysis is more thorough than van den Berg and Beintema (2000) or Beintema and van den Berg (2001).

Average time to contact (τ) is obtained more easily for a fronto-parallel plane than for a cloud-like stimulus since average time to contact is just the time to contact of the plane. Hence there is possibility that the findings of van den Berg and Beintema (2000) about precision of heading perception are specific to fronto-parallel planes. For more complex stimuli than a fronto-parallel plane, vector subtraction might be performed. Hence their arguments against vector subtraction do not seem definitive, though we adopted it for the present model. Further empirical studies may be needed for the definitive conclusion.

Although Beintema and van den Berg (2001) concluded that their experimental results supported their gain-fields-like model (Beintema and van den berg 1998). However, their results just imply models with vector-subtractive compensation for the rotational flow are implausible. Their gain fields model is not the only model that explains their data. The model presented in this paper explains the results, and other models without the vector subtraction may also explain them. A number of models have ways for making use of extraretinal information about self-rotation for heading computation other than vector subtraction. For template matching, only templates consistent with the extraretinal information may be used. The constraints about self-rotation can be easily combined with least-square methods. It would be interesting to perform simulations using the other models such as the template matching model of Perrone (1992).

6 General discussion

Based on Hanada and Ejima (2000a) model, we presented an algorithmic model that explains the data of van den Berg and Beintema (2000). The model does not use vector-subtractive compensation for the rotational flow. Instead, first the model computes the center of flow and estimates self-rotation, and then heading is recovered from the center of flow and the estimates of rotation. The new version of the model explains the results of the psychophysical experiments of van den Berg and Beintema (2000) at least qualitatively. Estimates of self-rotation derived from many sources can be processed identically for the present model, and fusion of different sources of information can be possible. We presented a fusion model of rotation estimates from various sources.

The algorithmic model explains human psychophysical data of heading perception very well (See Hanada and Ejima (2000a)). This indicates that explanations in the algorithmic level may be sufficient for many data. Thus, different implementations for the algorithm presented in this paper explain psychophysical data of heading perception to the same degree. The model presented in this paper is purely algorithmic. Also, we used only psychophysical data, and did not make use of physiological data for the modeling of heading computation in the brain. The model must be implemented by neural networks. Now we are developing a neural network that implements this algorithm. In near future, it will be presented. Physiological data will be compared to performance of the neural-network version of the model.

6.1 Coordinate transform

The transformed Z-axis (that is, the Z' axis) is passing through the center of flow. Selection of the other orthogonal two axes is free as long as the Z-axis is orthogonal to them. From a computational view of point, however, computation of heading recovery should be independent of the coordinate selection. However, the estimation of A' and B' is dependent on the coordinates in our model. A good way for achieving coordinate-independent computational results is to select the X' and Y' axes according to characteristics of the flow field like the selection of the Z' axis. For example, we can select the Y'-axis as the axis passing through a region in which the flow deviates most from the flow radiating from the center of flow after velocity components by rotation around the Z' axis are removed. For the model of human heading computation, however, the vertical and horizontal axes seem more natural. For animals on the earth, the directions would have special significances. Therefore we selected the vertical and horizontal directions for the directions of the X' and Y' axes. In practice, the dependence of computation on the coordinate selection is small and is not a critical problem for modeling of human visual functions.

A flow field is considered as a vector field (not necessarily continuous) on a 2-D closed manifold (a 2-D projective space). It may be possible that the present model is expressed in a more abstract form using concept of abstract vector fields on a Riemannian manifold without using any local coordinates. (For example, the center of flow was calculated by a method based on projective geometry presented by

Kanatani (1991) in the simulation above. It is one attempt for the calculation on a projective space.) For the concrete calculation on a manifold, however, certain local coordinates should be selected. We just selected local coordinates so that the algorithm could be written most easily. However, it should be noted that selection of local coordinates is not essential for the algorithm. The algorithm can be also written in the original local coordinates (where the line of sight is the Z-axis and the local neighbor is around the sight), or in the other coordinate systems (or possibly in an abstract form).

6.2 Static depth information

We pointed out that static depth information such as binocular disparities can be easily combined for the computation of the earlier model (Hanada and Ejima 2000a). It is also the case for the present model. Ratio of average time to contact to time to contact of each point is included in (17) and (18). The ratio is used for a substitute of ratio of average depth to depth of each point. When self-rotation occurs, τ and τ_i are rough approximations. If static depth information is available, the ratio of average depth to depth of each point can be directly computed. Thus, static depth information can be combined organically. van den Berg and Brenner (1994) reported that static depth information contributes to robustness of the heading judgement. (However, it does not reduce the bias in heading perception. See also Ehrlich et al. 1998)). We performed some simulations for our model varying the noise level, and found that robustness to noise increases with static depth information.

6.3 Visual estimation of heading

Under some conditions, human visual estimates of heading with self-rotation are accurate without extraretinal information, and visual compensation of self-rotation is nearly perfect. Although it has been not known what the sufficient and necessary conditions are for visual compensation for self-rotation, it has been suggested that the following conditions lead to nearly perfect visual compensation for rotation. 1) dense dots or textures (Li and Warren 2000). 2) sufficient motion parallax (Li and Warren 2000) 3) radial flow patterns whose center is located within the stimulus aperture (Lappe and Rauschecker 1994) 4) fixation to a static object (Warren and Hannon 1990). For the present model, these conditions also lead to good visual compensation for rotation. The model needs many sampling points to estimate rotation accurately because the model uses expectation values to estimate rotation. Also, the present model benefits from motion parallax because motion parallax highlights the deviation of the retinal flow from a radial pattern and the present model uses the deviation to estimate self-rotation in (16), (17) and (18) (See Hanada and Ejima 2000a). Moreover, the present model can estimate rotation accurately only when the center of low is located within the stimulus aperture because the sampling of velocities should be performed in a symmetric region around the center of flow for averaging in (16), (17) and (18). The center of flow is located within the stimulus aperture when an observer fixates a static object and moves. Furthermore, fixation to a static target keeps self-rotation fairly small. Hence the present model takes

advantage of fixation to a static target to estimate rotation. The present model can naturally explain the reasons why these conditions lead to good compensation for self-rotation.

6.4 Heading perception for a fronto-parallel plane

Grigo and Lappe (1999) showed that when ego-translation with ego-rotation toward a fronto-parallel plane is simulated without actual eye movement, observers perceived simulated heading for a large field of view (90 deg × 90 deg) and shortly presented flow. For a small field of view (60 deg × 60 deg) or for longer duration, however, incorrect heading is perceived, and the perceived heading is coincident with the center of flow. Our model cannot explain the data because the model cannot compute self-rotation for fronto-parallel planes from visual information alone (Hanada and Ejima 2000a).

However, our model may explain the results by two modifications: a method for estimating rotation for fronto-parallel planes was included in our model and reliance on the estimation method is reduced over time. Note that a new method for estimating rotation can be easily integrated because of the architecture of the present model in which rotation is estimated from various sources of information and the estimates are fused.

For a fronto-parallel plane, no vertical translation and no rotations around the vertical axis and the line of sight (i.e., $V = 0$ and $A = C = 0$), the velocity of a image point (x, y) is given by

$$u = \frac{-U + xW}{Z_f} - B - Bx^2 \tag{23}$$

$$v = \frac{yW}{Z_f} - Bxy \tag{24}$$

where Z_f is the depth of the fronto-parallel plane. W/Z_f and B is easily obtained from u of more than three points by using a linear least square method because u is linear about U/Z_f, W/Z_f and B. However, the estimate of B is not good. The contribution to u from the term B is indiscernible from that from the term U/Z_f, we must rely on the term Bx^2 to estimate B. However, since Bx^2 is much smaller than the other terms, the estimate of B is unreliable. However, we can also estimate B from the estimate of W/Z_f and a vertical velocity v. Since this method is relied on terms of Bx^2 in u and Bxy in v whose absolute values increases with eccentricity, the flow in the periphery visual field is needed. This explains why a large field of view is needed for correct perception of heading toward a fronto-parallel plane. (This method for recovering B above uses the horizontal and vertical velocity components separately. Instead, least square solution can be also obtained by a linear square method using both velocity components altogether since v as well as u is linear about U/Z_f, W/Z_f and B. However, both methods derive almost the same results.)

Decrease of reliance on this estimation method may explain correct perception for short duration. At the beginning of the stimulus, this visual estimate may be used for computing heading. The reliance of

this visual estimate may be reduced over time since this method is appropriate only to fronto-parallel planes. Perhaps the visual system relies more on (17) and (18) for the estimation of rotation.

It should be noted that the rotation estimation method may explain experimental data of Dyre and Andersen (1997). They showed that difference in the expansion rate for left and right visual fields affects heading perception, and induces perception of self-motion along a curved path. The proposed method for estimating rotation for a fronto-parallel plane relies on difference in vertical expansion rates at different visual areas caused by the term Bxy in v. For stimuli used by Dyre and Andersen (1997), perception of self-rotation may be induced by the difference in the vertical expansion rates. Perception of curved path reported by Dyre and Andersen (1997) is qualitatively consistent with perception predicted by the present model with the estimation method of rotation for fronto-parallel planes.

7 Conclusion

van den Berg and Beintema (2000) provided evidence against vector-subtractive compensation for the rotational flow. In this paper, an algorithmic model was presented. The model does not use vector-subtractive compensation for the rotational flow, and explains their data. Also, the present model explains many psychophysical data about heading perception (See also Hanada and Ejima 2000a).

References

1. Banks, M., Ehrlich, S. M., Backus, B. T., Crowell, J. A. (1996) Estimating heading during real and simulated eye movements. Vision Res 36:431-443
2. Beintema J. A., van den Berg A. V. (1998) Heading detection using motion templates and eye velocity gain fields. Vision Res 38:2155-79
3. Beintema J. A., van den Berg A. V. (2000) Perceived heading during simulated torsional eye movements. Vision Res 40:549-66
4. Beintema J. A., van den Berg A. V. (2001) Pursuit affects precision of perceived heading for small viewing apertures. Vision Res 41:2375-91
5. De Bruyn, B., Orban, G. A. (1988) Human velocity and direction discrimination measured with random dot patterns. Vision Res 28:1323-35
6. Crowell, J. A., Banks, M. S., Shenoy, K. V., Andersen, R. A. (1998) Visual self-motion perception during head turns. Nat Neurosci 1:732-737
7. Dyre, B. P., Andersen, G. J. (1997) Image velocity magnitudes and perception of heading. J Exp Psychol Hum Percept Perform 23:546-65
8. Grossberg, S., Mingolla, E., Pack, C. (1999) Neural model of motion processing and visual navigation by cortical area MST. Cereb Cortex 9:878-95
9. Ehrlich, S. M., Beck, D. M., Crowell, J. A., Freeman, T. C. A., Banks, M. S. (1998) Depth information and perceived self-motion during simulated gaze rotations. Vision Res 38:3129-3145
10. Grigo, A, Lappe, M. (1999) Dynamical use of different sources of information in heading judgments from retinal flow. J Opt Soc Am A 16:2079-91
11. Hanada, M., Ejima, Y. (2000a) A model of human heading judgement in forward motion. Vision Res 40:243-263
12. Hanada, M., Ejima, Y. (2000b) Method for recovery of heading from motion. J Opt Soc Am A 17:966-973
13. Hanada, M., Ejima, Y. (2000c) Effects of roll and pitch components in retinal flow on heading judgement. Vision Res 40:1827-1838
14. Heeger, D. J., Jepson, A. (1990) Visual perception of three-dimensional motion. Neural Comput 2: 129-137
15. Heeger, D. J., Jepson, A. (1992) Subspace methods for recovering rigid motion 1: Algorithm and implementation. Int J Comput Vis 7:95-117

16. Hildreth, E. C. (1992) Recovering heading for visually-guided navigation. Vision Res 32:1177-1192

17. Kanatani, K. (1991) Computational Projective Geometry. CVGIP Image understanding 54:333-348

18. Landy, M. S., Maloney, L. T., Johnston, E. B., Young, M. (1995) Measurement and modeling of depth cue combination: in defense of weak fusion. Vision Res 35:389-412

19. Lappe, M. (1998) A model of the combination of optic flow and extraretinal eye movement signals in primate extrastriate visual cortex. Neural model of self-motion from optic flow and extraretinal cues. Neural Netw 11:397-414

20. Lappe, M., Rauschecker, J. P. (1993) A neural network for the processing of optic flow from ego-motion in higher mammals. Neural Comput 5:74-391

21. Lappe, M., Rauschecker, J. P. (1994) Heading detection from optic flow. Nature 369:712-713

22. Lappe, M., Rauschecker, J. P. (1995) Motion anisotropies and heading detection. Biol Cybern 72:261-277

23. Li L., Warren W. H. (2000) Perception of heading during rotation: sufficiency of dense motion parallax and reference objects. Vision Res 40:3873-94

24. Longuet-Higgins, H. C., Prazdny, K. (1980) The interpretation of moving retinal images. Proc R Soc Lond B 208:385-397

25. Rieger, J. H., Lawton, D. T. (1985) Processing differential image motion. J Opt Soc Am A 2: 354-360

26. Royden, C. S., Crowell, J. A., Banks, M. S. (1994) Estimating heading during eye movements. Vision Res 34:3197-3214

27. Royden, C. S. (1994) Analysis of misperceived observer motion during simulated eye rotations. Vision Res 34:3215-3222

28. Royden, C. S. (1997) Mathematical analysis of motion-opponent mechanisms used in the determination of heading and depth. J Opt Soc Am A 14:2128-2143

29. Perrone, J. A. (1992) Model for the computation of self-motion in biological systems. J Opt Soc Am A 9:177-194

30. van den Berg, A. V. (1993) Perception of heading. Nature 365:497-498

31. van den Berg, A. V. (1996) Judgement of heading. Vision Res 36:2337-2350

32. van den Berg A. V., Beintema J. A. (2000) The mechanism of interaction between visual flow and eye velocity signals for heading perception. Neuron 26:747-52

33. van den Berg A. V., Beintema J. A., Frens, M. A. (2001) Heading and path percepts from visual flow and eye pursuit signals. Vision Res 41:3467-86

34. van den Berg, A. V., Brenner, E. (1994) Why two eyes are better than one for judgement of heading. Nature 371:700-702

35. Warren, W. H., Hannon, D. (1990) Eye movement and optical flow. J Opt Soc Am A 7:160-169

36. Young, M. J., Landy, M. S., Maloney, L. T. (1993) A perturbation analysis of depth perception from combinations of texture and motion cues. Vision Res 33:2685- 96

Figure captions

– Figure 1. Flow patterns due to ego-motion. (a) A flow pattern caused by translation. The center of flow corresponds to the heading point. (b) A flow pattern caused by rotation around the vertical axis. (c) A flow pattern caused by the translation for (a) + the rotation for (c). The center of flow shifts in the rotation direction. The center of flow no longer corresponds to the heading point.

– Figure 2. Coordinate system for a moving observer located at the origin. The observer moves with translation (U, V, W) and rotation (A, B, C). Point $P = (X, Y, Z)$ is projected to a point (x, y) on the image plane ($Z = 1$).

– Figure 3. The center of flow is the point that minimizes the square sum of the distance between the point and the line passing through the velocity flow vector.

– Figure 4. Transform of the original coordinate system to a new coordinate system so that the Z axis is passing through the center of flow.

– Figure 5. Compensation indices (CIs) predicted by the proposed fusion model and average CIs for human observers in Crowell et al. (1998) are shown. FBR: Full-body rotation, PHS: Passive head stabilization, AHS: Active head stabilization, PHP: Passive head pursuit, AHP: Active head pursuit, EP: Eye pursuit.

– Figure 6. For a lamellar flow, small noise disrupts the computation of the center of flow because the intersection points of some flow lines are located at the opposite of the actual center of flow.

– Figure 7. Simulation results for experiments of van den Berg and Beintema (2000). (a) SDs of heading estimated by the present model are shown as a function of the heading direction. (b) SDs of estimated heading are shown as a function of the retinal focus.

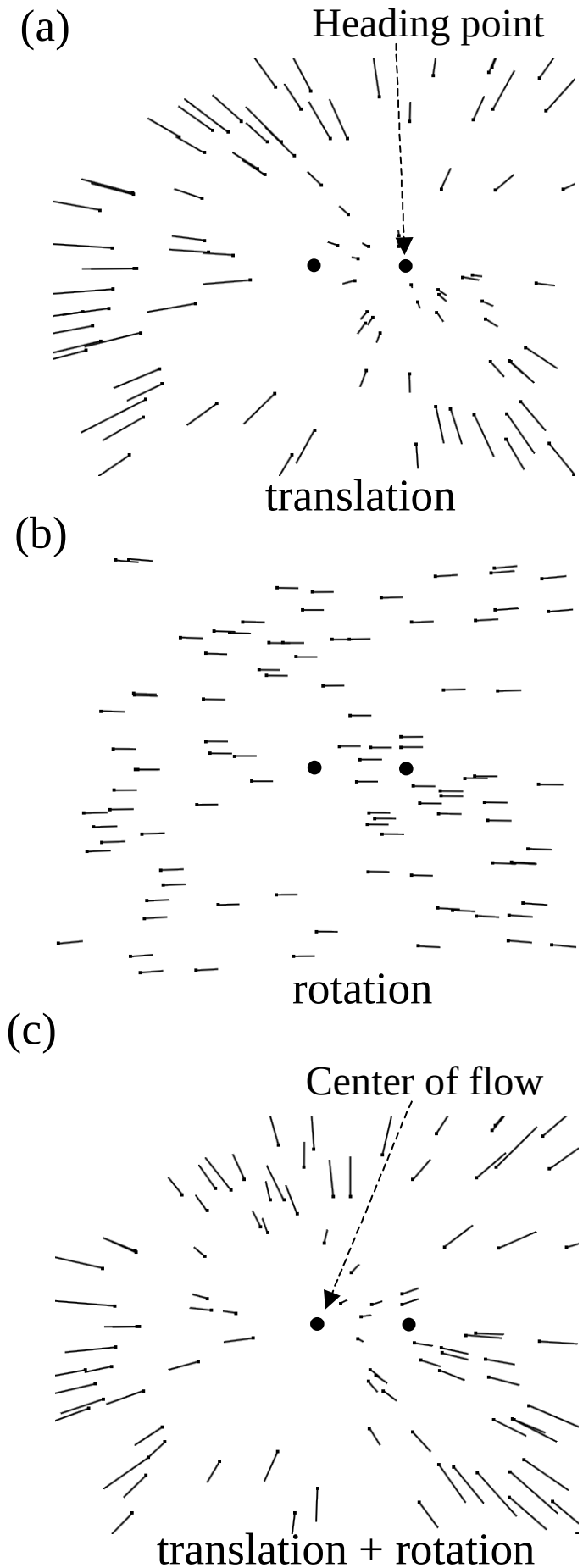
(a)
Heading point
translation
(b)
rotation
(c)
Center of flow
translation + rotation

Figure 1

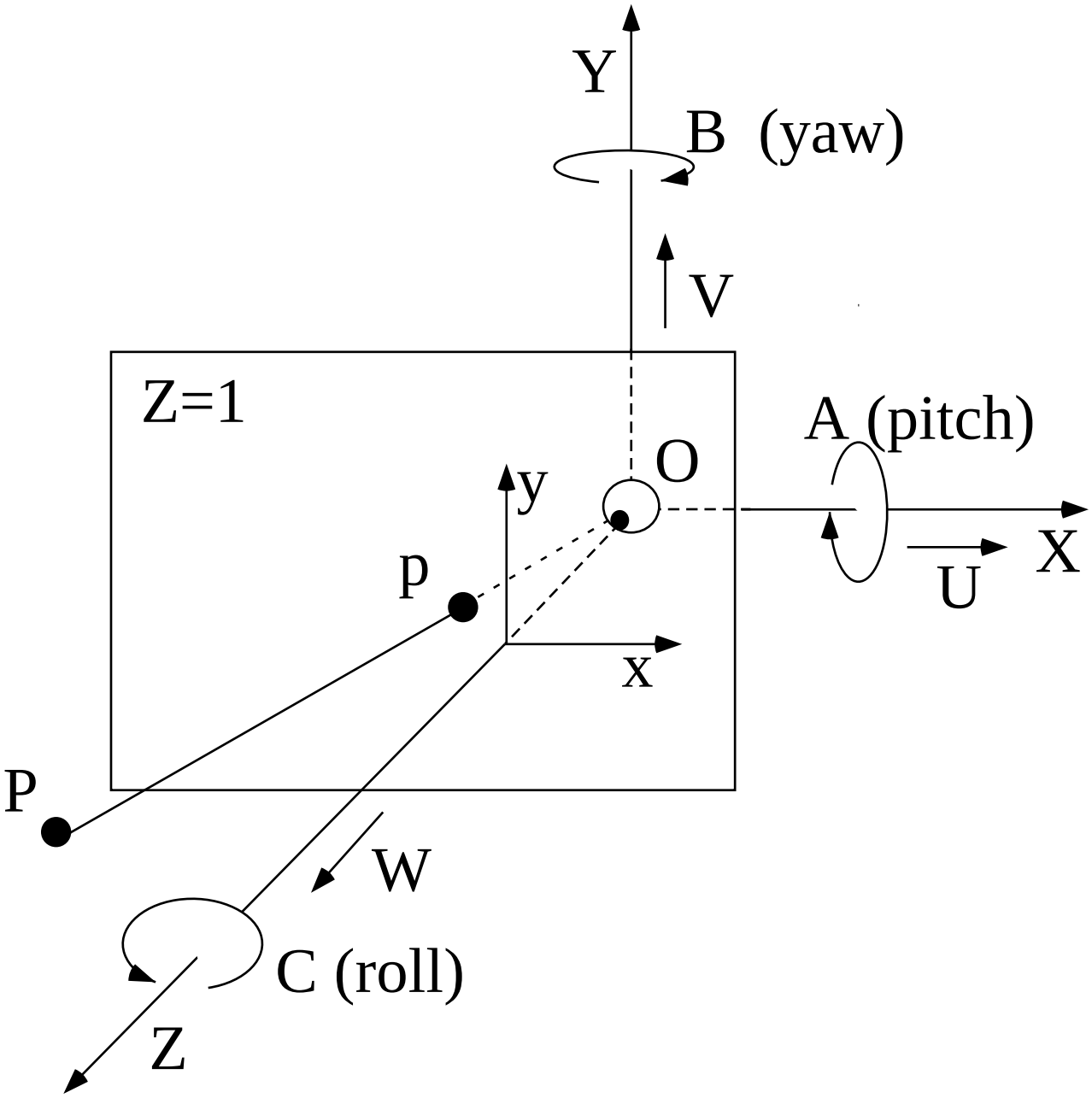

Figure 2

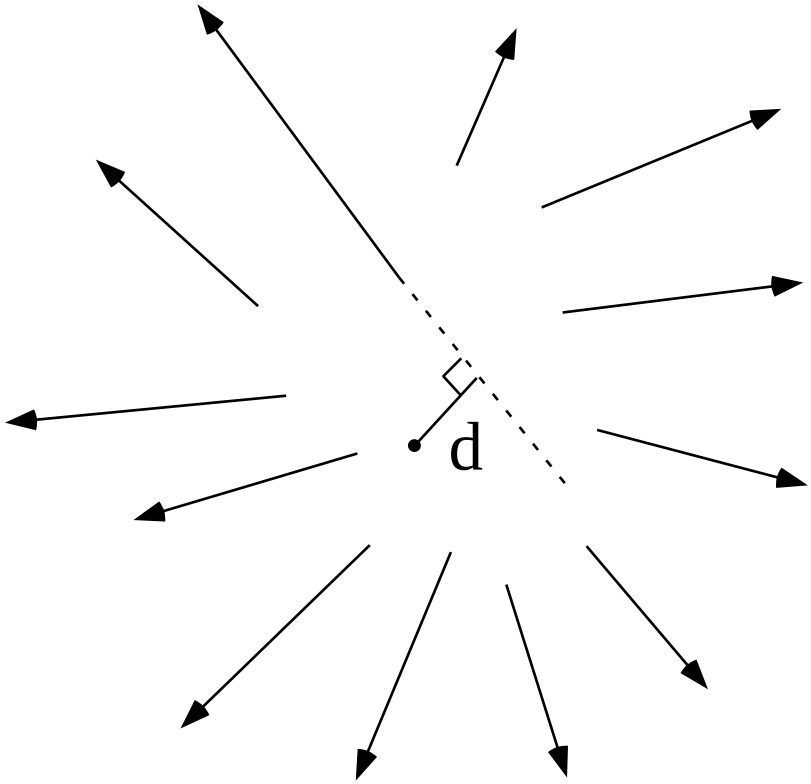

Figure 3

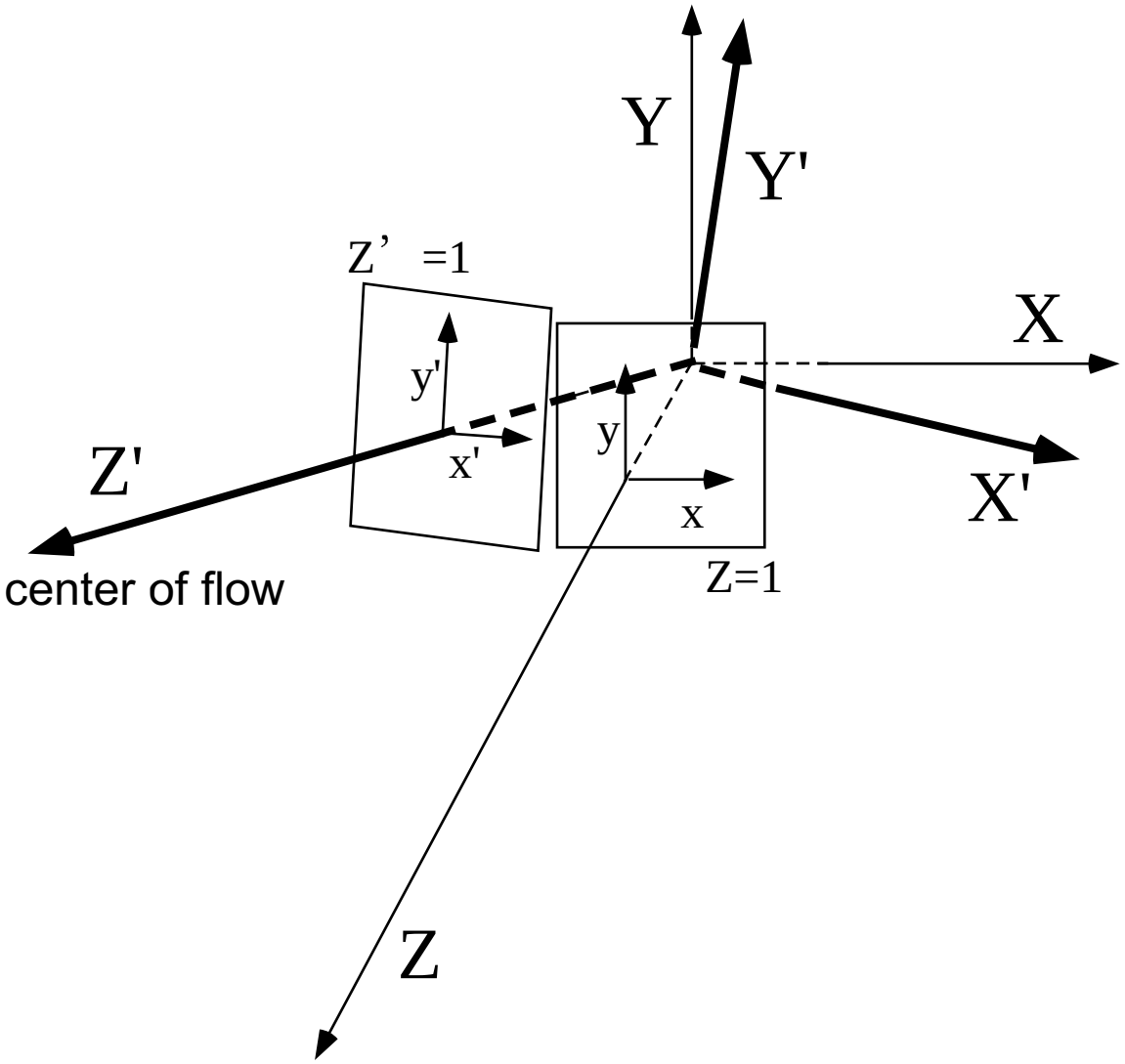

Figure 4

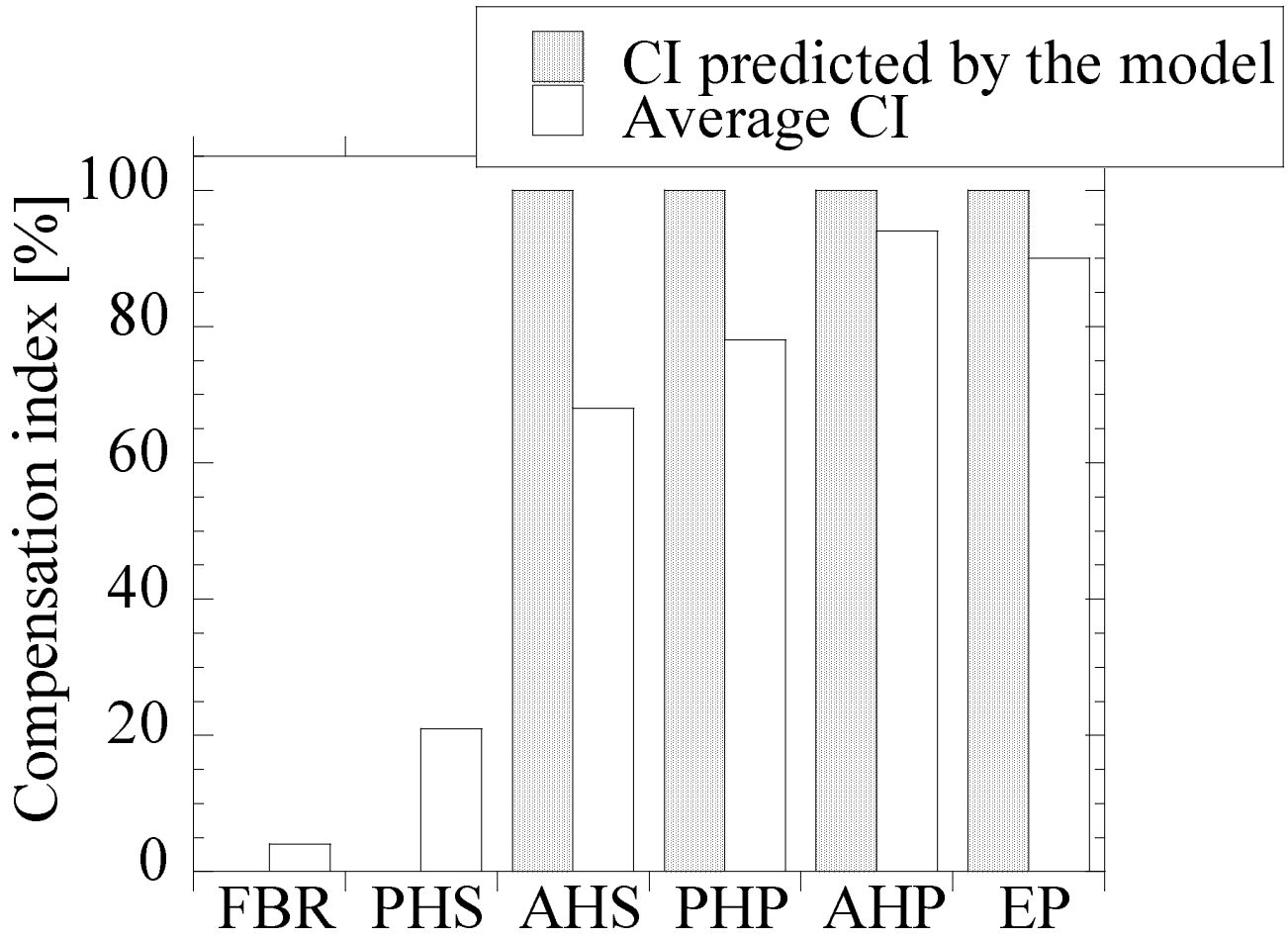
CI predicted by the model
Average CI
Compensation index [%]
100
80
60
40
20
0
FBR
PHS
AHS
PHP
AHP
EP

Figure 5

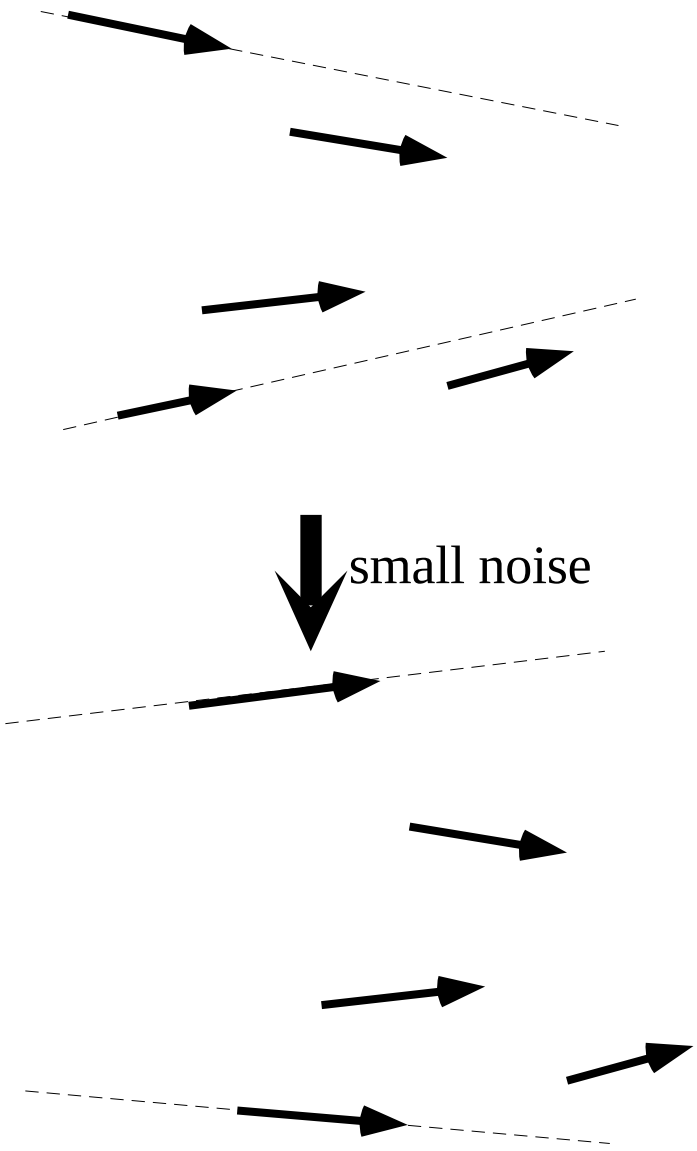

Figure 6

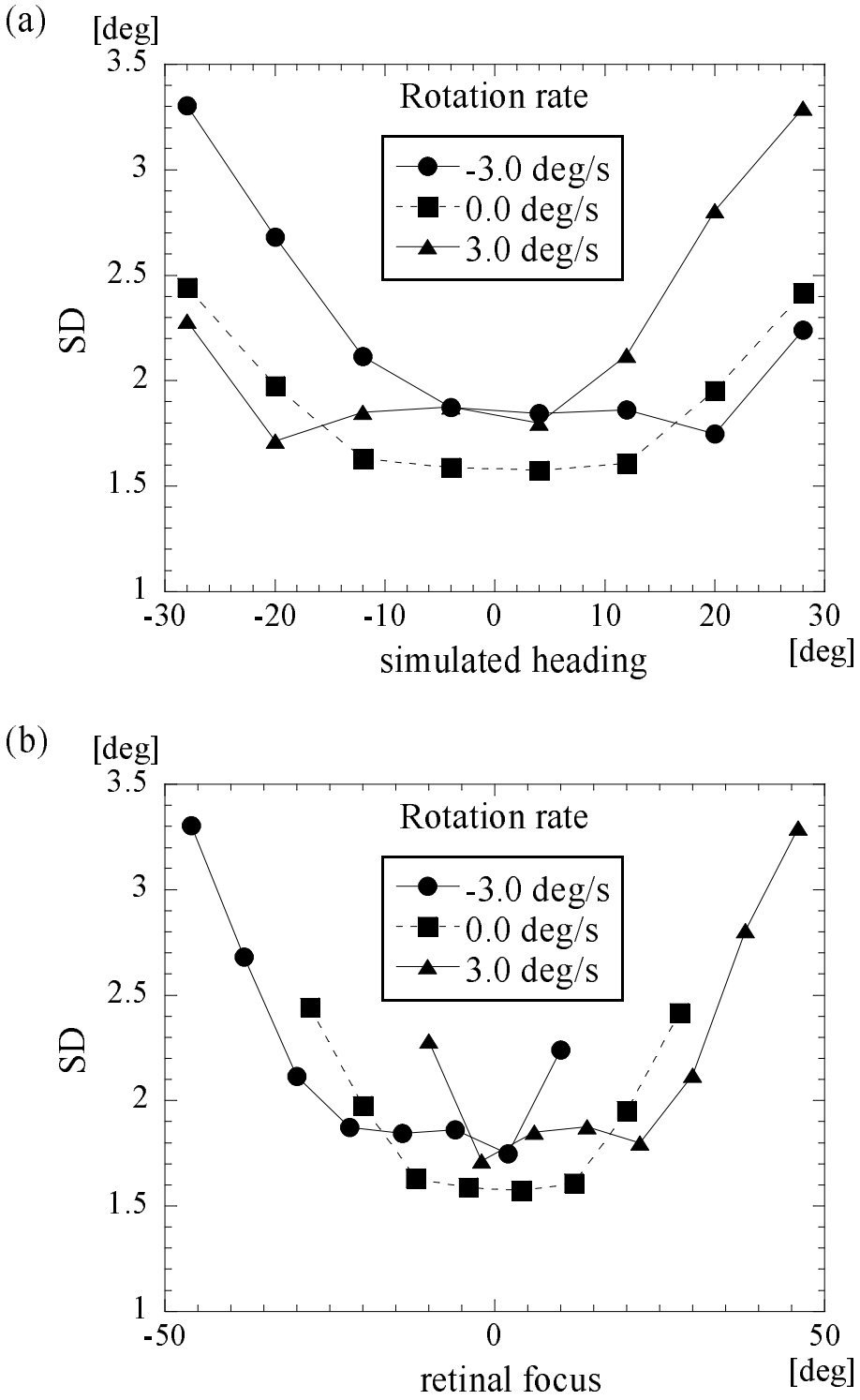
(a)
[deg]
3.5
3
2.5
2
1.5
1
SD
Rotation rate
-3.0 deg/s
0.0 deg/s
3.0 deg/s
-30
-20
-10
0
10
20
30
simulated heading
[deg]
(b)
[deg]
3.5
3
2.5
2
1.5
1
SD
Rotation rate
-3.0 deg/s
0.0 deg/s
3.0 deg/s
-50
0
50
retinal focus
[deg]

Figure 7